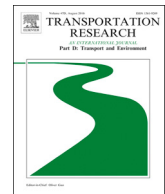




Contents lists available at ScienceDirect

Transportation Research Part D

journal homepage: www.elsevier.com/locate/trd

Development of a hybrid modelling approach for the generation of an urban on-road transportation emission inventory

Ran Tu, Islam Kamel, An Wang, Baher Abdulhai, Marianne Hatzopoulou*

Civil Engineering, University of Toronto, 35 St George St, Toronto, ON M5S 1A4, Canada

ARTICLE INFO

Keywords:

Traffic emission modelling
Greenhouse gas emissions
Traffic simulation
Microscopic emissions
Mesoscopic assignment
CLEVER

ABSTRACT

The development of accurate emission inventories at an urban scale is of utmost importance for cities in light of climate change commitments and the need to identify the emission reduction potential of various strategies. Emission inventories for on-road transportation are sensitive to the network models used to generate traffic activity data. For large networks (cities or regions), average-speed models have been relied upon extensively in research and practice, primarily due to their computational attractiveness. Nevertheless, these models are myopic to traffic states and driving cycles and therefore lack in accuracy. The aim of this study is to improve the quality of regional on-road emission inventories without resorting to computationally-intensive traffic microsimulation of an entire region. For this purpose, macroscopic, mesoscopic, and microscopic emission models are applied and compared, using average speed, average speed and its standard deviation, and instantaneous speeds. We also propose a hybrid approach called the CLuSTER-based Validated Emission Re-calculation (CLEVER), which bridges between the microscopic and mesoscopic approaches. CLEVER defines unsupervised traffic conditions using a combination of mesoscopic traffic characteristics for selected road segments, and identifies a representative emission factor (EF) for each condition based on the microscopic driving cycle of the sample. Regional emissions can then be estimated by classifying segments in the regional network into these conditions, and applying corresponding EFs. The results of the CLEVER method are compared with the results of microsimulation and of mesoscopic approaches revealing a robust methodology that improves the emission inventory while reducing computational burden.

1. Introduction

In Canada, transportation, the second largest greenhouse gas (GHG) emitter, produced 173.0 megatons of GHGs in 2015, according to Environment and Climate Change Canada ([Environment and Climate Change Canada, 2017](#)). In the city of Toronto, Canada's largest urban area, road transportation became the greatest GHG emitter in 2013, responsible for 41% of the total emissions, excluding rail, plane and boat ([Chief Corporate Officer, 2016](#)).

In order to quantify the impacts of travel demand and of traffic management on GHG emissions, various metropolitan areas around the world have developed road transportation emission inventories using a wide range of emission models. According to the level of detail, these models can be categorized as macroscopic or microscopic ([Rakha et al., 2003](#)). The required input for macroscopic models usually includes time-averaged speed, vehicle type, road type, and traffic condition ([Smit et al., 2010](#)). COPERT ([Borge et al., 2012](#)), EMFAC ([Park et al., 2011](#)), MOVES (in average speed mode) ([Koupal et al., 2003](#)) and MOBILE ([Fujita et al., 2012](#))

* Corresponding author.

E-mail addresses: ran.tu@mail.utoronto.ca (R. Tu), islam.kamel@mail.utoronto.ca (I. Kamel), ananan.wang@mail.utoronto.ca (A. Wang), baher.abdulhai@utoronto.ca (B. Abdulhai), marianne.hatzopoulou@utoronto.ca (M. Hatzopoulou).

<https://doi.org/10.1016/j.trd.2018.04.011>

generate emission factors (EFs) based on an average speed. In contrast, in the European model HBEFA (Colberg et al., 2005), EFs are determined based on different types of drive cycles (stop-and-go, free-flow, etc.). For microscopic models, their input includes detailed vehicle trajectories. For example, VERSIT+ (Ligterink and Lange, 2008, Smit et al., 2007) requires driving cycle variables like instantaneous speed, acceleration and road grades. Some microscopic models also calculate EFs from engine power or vehicle operating modes. PHEM (Hausberger et al., 2009), CMEM (Barth and Boriboonsomsin, 2008), and VT-MICRO (Rakha et al., 2004) are power-demand traffic emission models, in which the EF is a function of engine power and instantaneous speed. The United States Environmental Protection Agency (USEPA) model MOVES generates a distribution of vehicle operating modes based on second-by-second speed profiles, and assigns an EF for each operating mode to calculate total emissions.

Both macroscopic and microscopic approaches to model emissions have been widely applied in previous studies, and their advantages and shortcomings have been discussed. Macroscopic approaches are time-efficient, with minimal computational needs (Jamshidnejad et al., 2017) and the input data are relatively easy to retrieve from traffic measurements or modelling, while ignoring the variation in driving behaviours (Rakha et al., 2011). Microscopic approaches consider instantaneous speeds, but are more time-consuming (Rakha et al., 2011), especially for large networks. Recently, mesoscopic approaches have begun to gain momentum, capitalizing on the advances in traffic assignment models, such as dynamic traffic assignment. Mesoscopic traffic models generate average traffic speeds but also include number of stops, delay, or queue lengths. Using the output of mesoscopic traffic assignment models, various studies have proposed to construct synthetic drive cycles from average speeds, number of stops, and stop delay, while dividing trips into various categories based on driving mode. EFs are then applied for each mode to calculate total emissions. For example, Gori et al. proposed a mesoscopic approach which divided links into three parts: the length in which vehicles were at (1) free-flow speed, (2) in the queue, and (3) in acceleration phase, and applied an emission rate to each portion (Gori et al., 2013). VT-Meso developed by Rakha et al. constructed a synthetic drive cycle based on the number of vehicle stops, stopped delay, and average speed. The synthetic drive cycles were then utilized to generate average emission rates based on the VT-MICRO model, and total emissions were calculated by multiplying emission rates with vehicle kilometers travelled (Rakha et al., 2011). Jamshidnejad et al. (2017) developed another mesoscopic approach whereby traffic conditions were divided into three states: under-saturated, saturated, and over-saturated. Within each state, trips were categorized into four groups according to average speed and number of stops. A representative drive cycle for each group was selected and emissions were then calculated based on a microscopic emission model.

While the proposed mesoscopic approaches provide more sensitivity to traffic conditions than average speed models, they still fall short of providing emission estimates that are close to the output of microscopic models, which are capable of incorporating the full drive-cycle. In this paper, three scopes of traffic simulation and emission modelling are presented. On the one end, a full traffic microsimulation was conducted, and GHG (in CO_{2eq}) emissions were estimated based on instantaneous speeds. On the other end, average speeds from mesoscopic simulation were used to generate EFs for estimating emissions. Besides, mean and standard deviation of speed for each road section from mesoscopic simulation results were expanded to a distribution; a weighted EF was then generated. Finally, we propose an approach which bridges between the microscopic and mesoscopic models. The CLustEr-based Validated Emission Re-calculation (CLEVER) method relies on a microsimulation of a small sample of roads in a large road network in order to develop EFs, which are representative of specific traffic conditions categorized by mesoscopic traffic data across the network. The results of the CLEVER method were compared against the results of microsimulation and of mesoscopic modelling, revealing a robust approach, which improved the emission inventory while reducing the computational burden of microsimulation.

2. Methodology

2.1. Study area and data

This paper used the road network in downtown Toronto (Fig. 3a). Our study domain covered an area of 3.31 km from North to South, and 4.53 km from West to East. The total road length in both directions is 165.8 km. The network is composed of around 79%

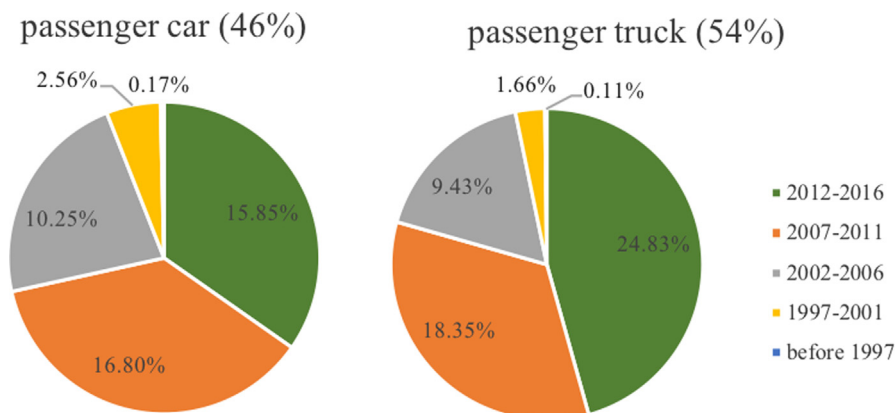


Fig. 1. Distribution of vehicle types and model years.

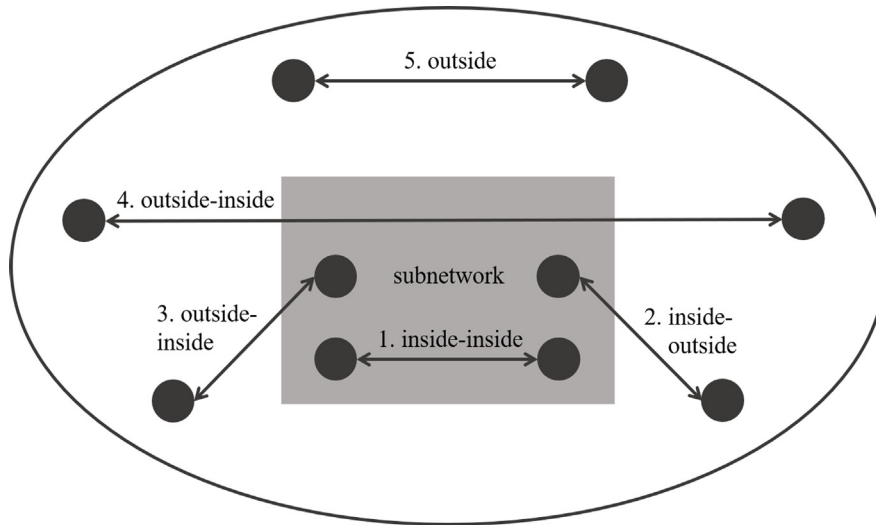


Fig. 2. Relationship between trips and sub-network.

arterials, and 21% freeways. On-road vehicle composition was derived from the 2016 Statistics and Management Reporting of the Ontario Ministry of Transportation (Ministry of Transportation Ontario, 2017), which is shown in Fig. 1.

Microsimulation of traffic and emissions was generated for the downtown network. The same network was then used to generate a mesoscopic traffic assignment, with shorter computational time than microsimulation, generating for every road link an average speed and the standard deviation of speed. These two parameters were in-turn used to compute emissions, which were then compared with the emissions estimated using microsimulation. Finally, the CLEVER method was introduced to bridge the gap between the micro and meso approaches. The CLEVER method uses the output of the mesoscopic simulation to identify clusters of road segments based on their traffic conditions. Each cluster was assigned a representative EF and was associated with mesoscopic traffic variables. The outcome is an approach that associates mesoscopic variables with representative EFs (calculated and validated based on microsimulation).

Traffic assignment was conducted using the platform AIMSUN, which allows for traffic microsimulation as well as mesoscopic traffic assignment. The Transportation Tomorrow Survey (TTS) 2011 AM peak OD matrix was used as the network traffic demand. There were approximately 1.7 million trips between 6:00 am and 10:00 am travelling through the entire Greater Toronto Area. To obtain the OD matrix for the downtown area, the ‘OD traversal’ function in AIMSUN was applied. This function summarizes the total amount of trips that cross the study boundary (in-out, out-in, out-out), or happen inside the boundary (in-in) (TSS-Transportation Simulation Systems, 2017), as illustrated in Fig. 2. The total demand of the downtown was around 72,000, and the simulation time was extended to 12:00 pm to ensure that all the trips can be completed.

After demonstrating the robustness of the CLEVER approach in the downtown network, we presented an application to the whole city of Toronto (Fig. 3b). The road network in the city included more than 15,000 road links, with about 654 km highway segments, and 2328 km arterials. In the city network, a sample of roads (242.6 km) was selected and used for traffic microsimulation in order to derive cluster-representative EFs. The full city network was simulated in mesoscopic scale, with microscopic simulation only in the sample area. The vehicle composition adopted in this case was the same as the one used for the downtown network (Fig. 3a). Similarly, traffic demand for the city network was also derived from the trip diary data by applying ‘OD traversal’ in AIMSUN. In total, there were about 490,000 trips in the city network.

2.2. Calibration and validation of traffic model

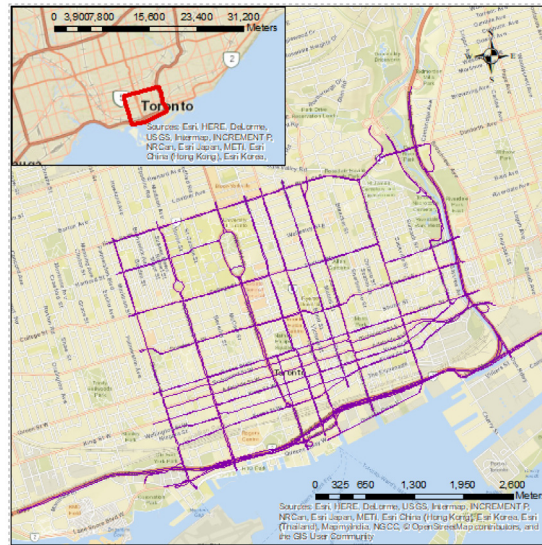
Two types of on-road data were adopted for model calibration. First, loop detector data from 176 locations on major highways of the region were collected every 20 s. They were aggregated to hourly average speeds and traffic volumes to compare with the average speed and traffic volumes derived from AIMSUN mesoscopic simulation. Several network parameters, such as road capacity, jam density, traffic demand, and traffic signals, were then adjusted to minimize the main calibration criterion GEH, which is defined as:

$$GEH = \sqrt{\frac{2(M-C)^2}{M+C}} \quad (1)$$

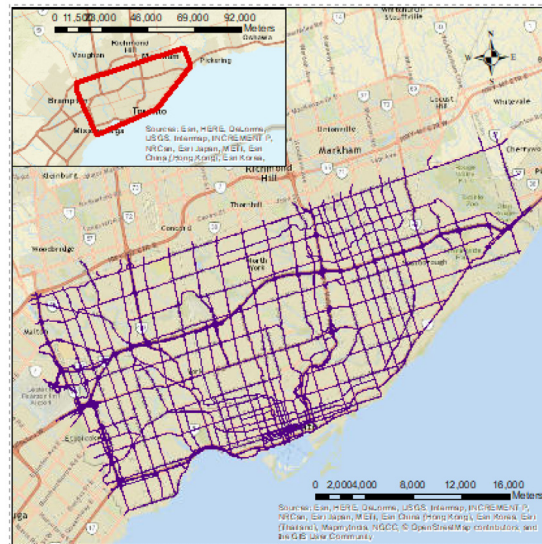
where:

M is simulated traffic flow, and C is observed traffic flow.

Fig. 4 shows the calibrated speeds and volumes, and as a result, the average GEH after calibration was 13.6 (down from 20). More



a. Downtown network for CLEVER development and comparison



b. City of Toronto network for CLEVER application

Fig. 3. Study area for model development (a) and model application (b).

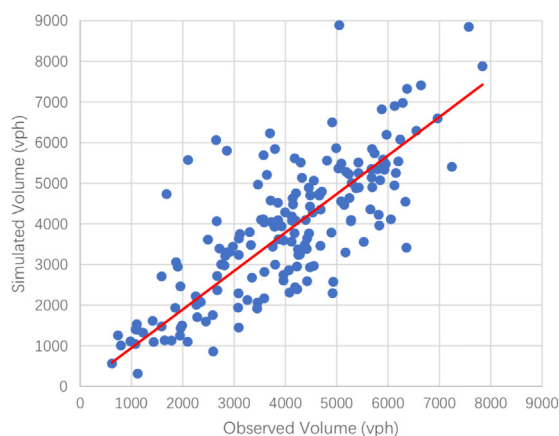
detail on the network calibration can be found in [Kamel et al. \(2018\)](#).

In addition, instantaneous speeds of passing vehicles recorded by a roadside radar on College St, a main downtown artery, were examined. Second-by-second speeds at the same spot as the radar location in the AIMSUN network were selected from the trajectory profile. Speed distributions from observed (radar) and simulated (AIMSUN) traffic are illustrated in [Fig. 5](#). The figure illustrates that the speed variation from AIMSUN was smaller than the radar speed, while the average values were similar (no significant difference at 95% confidence).

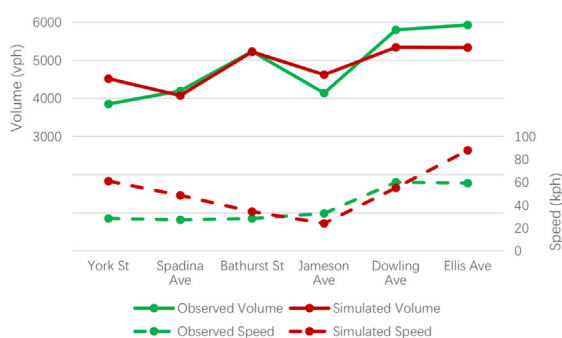
2.3. Traffic modelling

In this paper, the AIMSUN software platform was used for traffic simulation. It is capable of assigning traffic at four different levels: macroscopic traffic assignment, mesoscopic, microscopic, and hybrid. In this paper, mesoscopic, microscopic and hybrid simulation were applied. The inputs utilized in the three scales were OD matrices derived from 2011 TTS data (as described above).

In AIMSUN, the process of microscopic simulation is called “time-based”. Time moves according to timesteps and intervening



a. Observed vs. simulated traffic volumes



b. Observed vs. simulated volumes and speeds over the Gardiner expressway

Fig. 4. Model validation results.

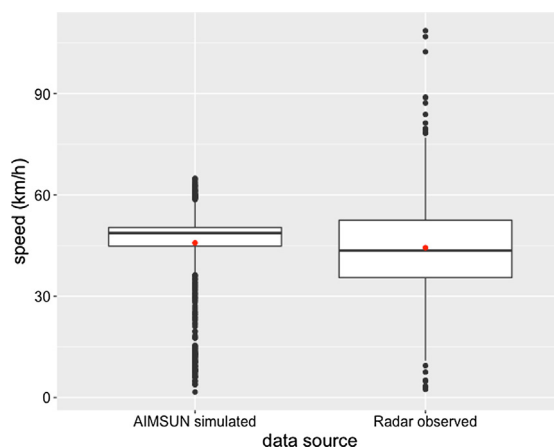


Fig. 5. Observed vs. simulated speed distributions on College Street (the horizontal line inside the box is the median value of speeds, and the red dot represents the mean value).

movements. For example, speed and lane changing are considered for each vehicle, while the vehicle moves by the distance determined for that timestep. In this way, the vehicle instantaneous speed and acceleration of each timestep (one second) can be recorded by microscopic simulation. However, mesoscopic simulation is discrete and “event-based”; time moves forward to the next “event”, which marks a vehicle entering or exiting a link. The trajectory of vehicles only includes these events, while the activities in between are ignored. Mesoscopic simulation generates average traffic conditions within the simulated time interval, and standard deviation of each traffic variable is also calculated based on each event during this interval (TSS - [Transport Simulation System](#),

2017), therefore the record of mesosimulation is based on each link. In this paper, the time interval of mesoscopic simulation was one minute.

The hybrid simulation can be regarded as running mesoscopic assignment and microsimulation at the same time. First, a microsimulation area is selected as a sample within the network. When running hybrid simulation, the whole network is simulated by a mesoscopic model; at the same time, the selected area is simulated by a microscopic model. In this process, the mesoscopic and microscopic simulations are synchronized, exchanging information on vehicles. The output of hybrid simulation includes a mesoscopic result for the whole network and microscopic result for the microsimulation area (TSS - Transport Simulation Systems, 2017). The time interval of the mesosimulation portion of hybrid simulation in this paper was set at 1 h.

2.4. Emission modelling

The MOVES model developed by the USEPA was used for the generation of EFs. MOVES estimates emissions by determining the vehicle operating mode, which is based on the vehicle specific power (VSP) as illustrated in Eq. (2) (United States Environmental Protection Agency, 2015).

$$P_{V,t} = \frac{Av_t + Bv_t^2 + Cv_t^3 + mv_t a_t}{m} \quad (2)$$

where:

$P_{V,t}$ is the vehicle specific power (VSP);

v_t is the speed at time t (m/s);

a_t is the acceleration at time t (m/s²);

m is the mass of the vehicle, usually referred as “weight” (Mg);

A, B and C are track-road coefficients, representing rolling resistance, rotational resistance and aerodynamic drag, in unit kW-sec/m, kW-sec²/m² and kW-sec³/m³.

Two formats for vehicle activity inputs are available to use in MOVES. First, by providing second-by-second speed profiles, MOVES calculates an operating mode (opMode) distribution of the driving cycle. Combining with meteorology, fuel type and vehicle type, MOVES generates an emission rate for each opMode, then generates total emissions using emission rates and the opMode distribution. In this paper, this mode was applied to estimate emissions based on microscopic traffic simulation (Fig. 6a).

Second, by providing average speed as well as road characteristics, MOVES assigns a default driving cycle that best fits the given data, and provides an EF associated with each average speed. Total emissions are calculated by multiplying EFs with vehicle kilometers travelled (VKT). This approach is referred to as “average-speed” in this paper (Fig. 6a).

Besides average speed $\mu(v_t)$, the mesoscopic simulation in AIMSUN also returns standard deviation of speed $\sigma(v_t)$, which indicates the variation of the event average speed during the simulated time interval. To make use of this information, we assumed that speeds on each link followed a normal distribution $N(\mu(v_t), \sigma(v_t))$ (Iannone et al., 2013). A weighted EF was calculated based on a normal distribution, which can be illustrated in Eq. (3). This approach was referred to as “meso” in this paper (Fig. 6a).

$$\overline{EF}_t = \int_{\mu(v_t)-2\sigma(v_t)}^{\mu(v_t)+2\sigma(v_t)} P(v_{it}) * EF(v_{it}) \quad (3)$$

where:

\overline{EF}_t is weighted EF of time t ;

$\mu(v_t)$ is average speed from mesoscopic result;

$\sigma(v_t)$ is standard deviation of speed from mesoscopic result;

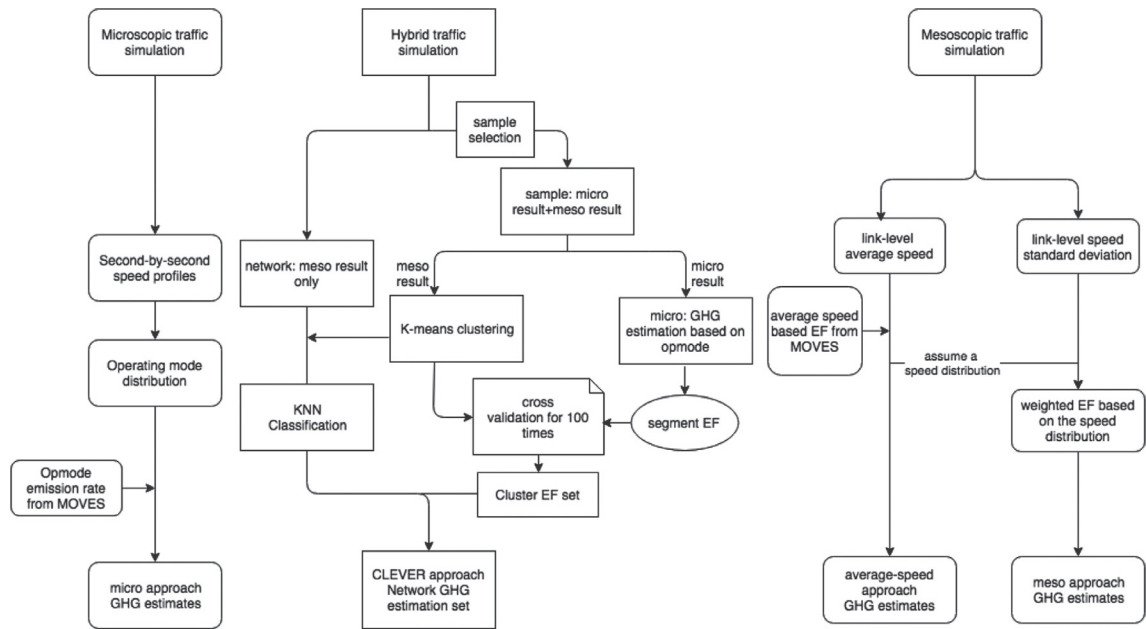
$P(v_{it})$ is probability density of v_{it} which follows normal distribution

$N(\mu(v_t), \sigma(v_t))$. The domain of v_{it} is $(\mu(v_t)-2\sigma(v_t), \mu(v_t) + 2\sigma(v_t))$;

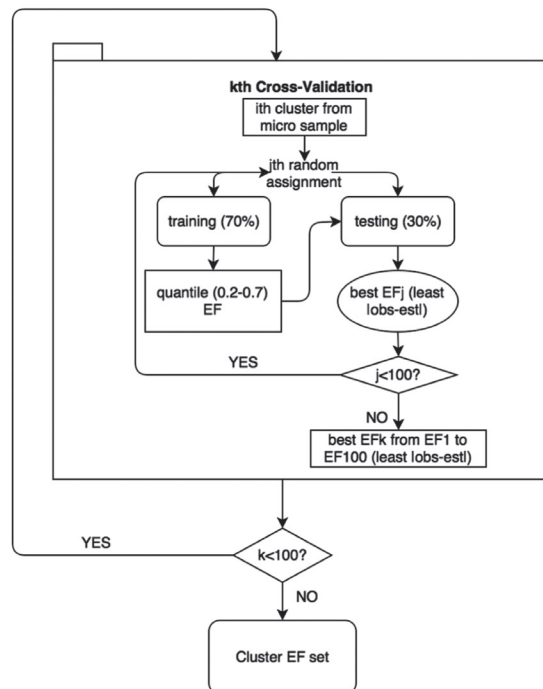
$EF(v_{it})$ is EF of v_{it} derived from MOVES.

Two EF matrices based on average speed and opMode ID were derived from MOVES separately. The estimation of total emissions was generated outside of the MOVES-GUI using the EF matrices: the average speed-based EF matrix was applied to the average-speed and meso approach, while the opMode-based emission rate matrix was used for microscopic traffic data. The extraction process of EF matrices was similar to the one published by Liu et al. (2016). Liu et al. demonstrated that this method can largely improve the computational speed compared to the MOVES-GUI (Liu et al., 2016).

To keep the on-road information consistent with the traffic demand in this paper, MOVES day/time was defined as the morning of a weekday in June. The temperature was set at 60F, humidity being 70, based on the average meteorology conditions of June in Toronto. Since the fleet composition data from MTO was for gasoline passenger vehicles, the fuel vehicle combination in MOVES was set as gasoline-passenger car and gasoline-passenger truck (source type 21 and 31).



a. Algorithms of emission models using different scopes of traffic simulation



b. Cross-validation approach in the CLEVER method

Fig. 6. Approaches for emission modelling across various scopes.

2.5. CLEVER approach

We propose a new approach for GHG estimation using hybrid traffic simulation. The algorithm follows the structure of clustering-validation-classification, so it is referred to as the CLustEr-based Validated Emission Re-calculation (CLEVER). The algorithm of CLEVER is displayed in Fig. 6a and referred to as “hybrid traffic simulation”.

This approach groups road segments based on their hourly average traffic conditions, and assigns a representative EF for each

group. However, the categories are not pre-defined, and the categorization is not only based on average speed and stops. Instead, segments are grouped into different clusters through an unsupervised clustering k-means method.

To start, a random sample of road segments within the study area is selected. To ensure representativeness of the sample, selected links should include all road types, and the major traffic characteristics in the sample should be representative of the network. Using AIMSUN, a hybrid traffic simulation is performed, which means that the sample of links will undergo microsimulation while the entire network will undergo mesoscopic assignment.

For segments in the sample, using the MOVES EF Matrix (opMode estimation mode), GHG emissions (in CO_{2eq}), and the EF of each segment are calculated. Then, using the k-means clustering method, these segments are grouped into 30 clusters based on selected traffic variables derived from the mesoscopic assignment. In this paper, besides average delay and speed which have been commonly used in previous research, standard deviation of these two variables, link capacity (an important indicator for the road type), link density (which is vital to evaluate the level of service), and vehicle kilometers travelled were also used as the criteria for clustering. Once the segments within the sample are tagged to different clusters, a center of each cluster is generated in the form of the feature vector.

Subsequently, a representative EF for each cluster is generated using a cross-validation approach. The cross-validation process is illustrated in Fig. 6b. In each cluster, segments in the sample are randomly assigned as training set and testing set following the ratio of 70 (training): 30 (testing). Candidate EFs are selected from EFs of segments in the training set. Instead of automatically selecting the average EF in each cluster as the representative EF for the cluster, this approach does the following: the algorithm first chooses the EF of the 20th percentile in the training set, and uses this EF to calculate the total emissions for segments in the testing set. Then the error between the emissions calculated using the MOVES EF Matrix (which is regarded as the baseline) and the emissions calculated using the selected EF for the cluster, is computed for segments in the testing set. Next, the algorithm chooses the EF of the 22nd percentile, and goes through the same procedure. The process repeats by adding 2% each time until the 70th percentile is reached, therefore 26 EFs are listed. The representative EF for each cluster is the one that minimizes the error in the testing set. Since the training set is selected randomly from the sample, this procedure is repeated 100 times, for various randomly assigned training-testing sets, always in the ratio 70:30. For each cluster, the EF that leads to the smallest error in the testing set, after 100 iterations, is selected as the EF representing that cluster of segments. In order to examine the variability in the chosen EFs, the algorithm also repeats the cross-validation procedure described above 100 times and stores 100 EFs, each representing the EF that has led to the smallest error within each cross-validation procedure. The selected EF is associated with the feature vector formed with mesoscopic traffic simulation variables (average speed, delay, etc.) and these become the mesoscopic variables describing the cluster.

To calculate network GHG, the remaining road segments are assigned to the different clusters with the use of K-Nearest Neighbors (KNN) algorithm, based on the variables generated by the mesoscopic assignment (Eq. (4)). Emissions of each segment are calculated based on the representative EF for the cluster.

$$D_{ik} = \sum_{j=1}^c (V_{k,j} - V_{i,j})^2 \quad (4)$$

where:

- D_{ik} is the distance from the segment k to the center of cluster i ;
- $V_{k,j}$ is the traffic variable j of segment k ;
- $V_{i,j}$ is the traffic variable j of the center of cluster i ;
- c is the dimension of the feature vector (seven in this paper).

3. Results

The results for the downtown network were presented first. In this case, results from the micro, average-speed, and meso (average and standard deviation of speed) approaches were illustrated. Then the results of the CLEVER approach were presented for the downtown network and compared with results from micro, average-speed, and meso approaches. After demonstrating the robustness of CLEVER in the downtown network, the results from average-speed and CLEVER approaches for the entire city of Toronto network were presented and compared. Results of micro approach for the entire city network were not available due to high computational burden from full-network microsimulation.

3.1. Comparison between micro and meso GHG estimates

The total GHGs emitted in the downtown network were estimated using three approaches: (1) the mesoscopic approach with road average speed (average-speed), (2) mesoscopic approach with average speed and standard deviation of speed (meso), and (3) full microsimulation (micro). The results are presented in Fig. 7.

Based on the micro approach, the network produced 70.06 tons of GHGs. The average-speed approach led to the greatest underestimation of GHG emissions, which was 30% lower than the result obtained using microsimulation. This was due to the fact that this approach did not include the effects of idling, acceleration and deceleration. With consideration of the standard deviation of mean speed, the meso approach resulted in a slightly higher emission estimate compared to average-speed. However, it remained significantly lower than the estimate based on the microsimulation approach. This result indicated that traffic conditions were not

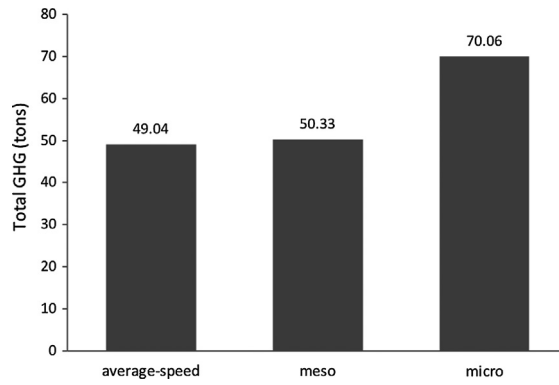
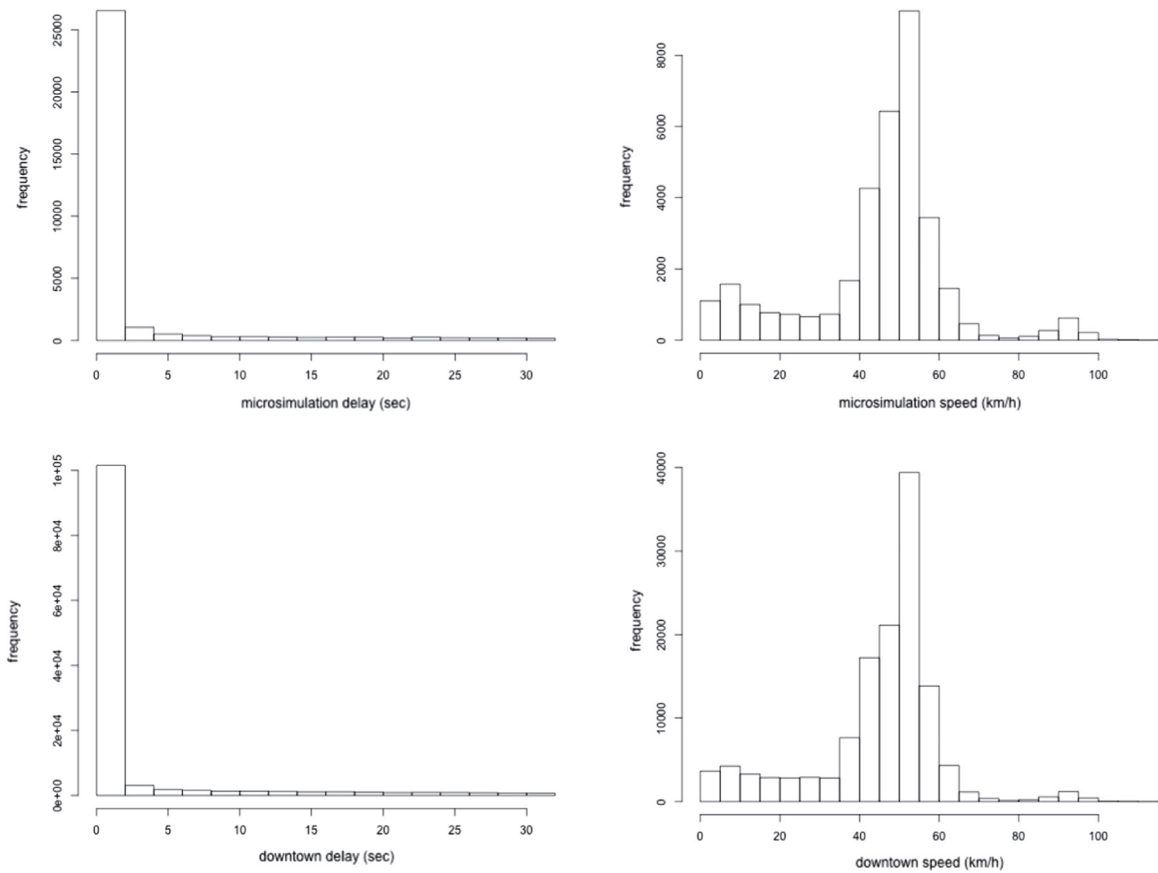


Fig. 7. Comparison between GHG emission estimates for the downtown network (average-speed is the method using only average speed; meso is the method considering both average speed and standard deviation of speed; micro indicates microscopic approach).



a. Delay distribution for sample segments (top) and full downtown (bottom)

b. Speed distribution for sample segments (top) and full downtown (bottom)

Fig. 8. Delay, speed, density and flow distribution of sample area and of downtown road network.

homogeneous for the road links across the network, therefore the speeds of all events during the time interval in mesoscopic simulation cannot be assumed to follow a normal distribution. Previous research has shown that under congestion, speed distributions across each link will follow other distributions such as bi-modal (Jun, 2010).

3.2. Emission estimate using the CLEVER method

Using the same downtown network, we extracted a random sample of road segments while making sure that different types of

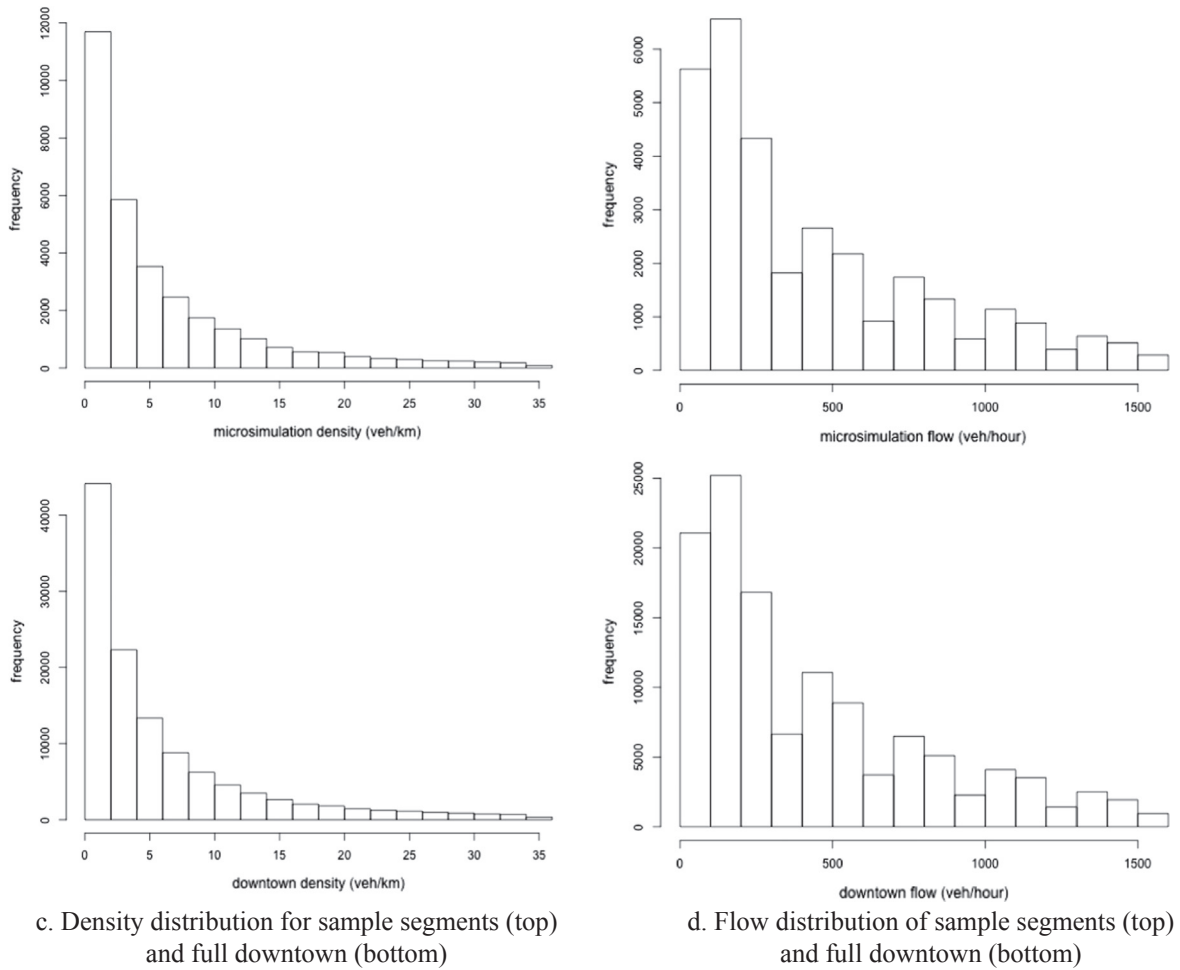


Fig. 8. (continued)

roads were represented. The sample microsimulation area in the downtown consisted of 39.3 km of arterial roads, and 14.6 km of highways. The ratio of arterials over the microsimulation area was about 73%, which was close to the ratio of arterials in the full downtown network (79%).

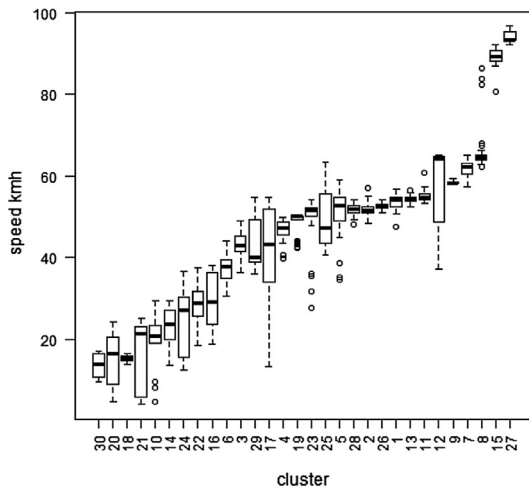
Since clusters generated by the CLEVER method are supposed to reveal potential on-road traffic conditions, traffic characteristics of selected samples should be representative of the entire network. Fig. 8 shows the distribution of delay, density, speed and flow in the sample area and the downtown network. The selected samples showed similar traffic distribution as the entire network.

The sample segments were clustered based on the k-means approach into 30 different clusters. Fig. 9a illustrates the average speed distribution within each of the 30 clusters. Some of the clusters exhibited a high variability in average speeds across the different segments within the cluster while other clusters had a more homogenous distribution among segments. This shows that besides average speed, other variables may also be important for clustering, stressing the need to cluster based on a combination of variables.

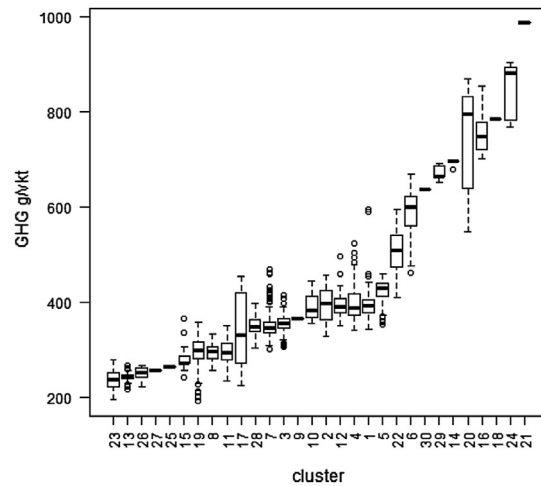
Fig. 9b illustrates GHG EFs of each cluster obtained after applying the CLEVER approach, including repeating 100 times the cross-validation process. We observed that around 30% of clusters had EFs higher than 600 g/VKT (corresponding to an average speed of around 20 km/h in the average-speed approach), possibly showing saturated or over-saturated traffic conditions represented by these clusters. Another possibility was that the segments in these clusters were close to signalized intersections thus including a lot of idling, accelerating and decelerating activities. Validated EFs of most clusters had limited variability, while a small number of clusters had a relatively greater range of EFs, indicating a greater stochastic influence of the cross-validation process for these clusters. This may be due to a higher variation of segments' EFs within these clusters.

Fig. 9c presents the estimation errors (compared to the results from the micro approach as the baseline) of the testing sets (the 30% random sample in each cluster). After 100 repetitions, most of the clusters had errors distributed around 0, representing high accuracy and representativeness of validated EFs. Due to the small number of segments included in clusters 18, 25, 27 and 30, which had 3, 4, 5, and 4 segments respectively, the EFs of segments within these clusters were not homogeneous, and errors of these clusters were not centered at 0.

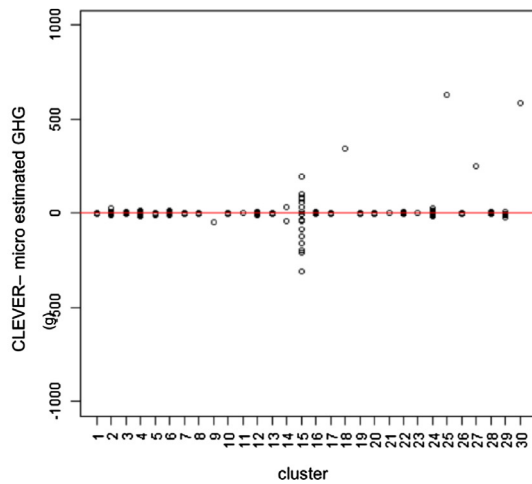
Fig. 9d presents the CLEVER result with 100 repetitions. The estimated GHG emissions for the network ranged between 67.4 tons



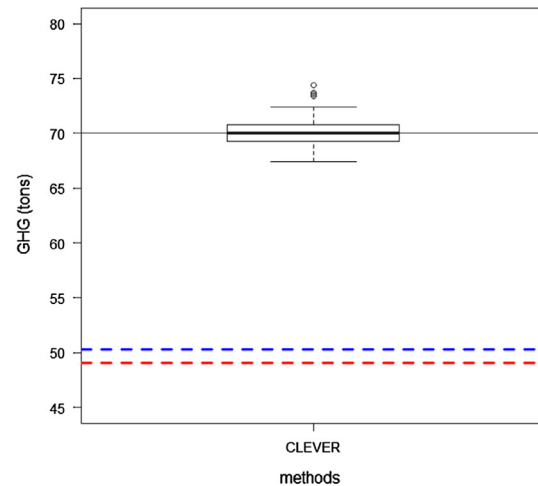
a. Speed distributions across each of the 30 clusters (sorted by mean speed value of each cluster)



b. Distribution of representative EFs based on the CLEVER method; in each cluster, 100 values were generated based on the 100 random samples for the testing set (sorted by mean EF value of each cluster)



c. CLEVER minus micro emissions in the testing sets, based on 100 random samples of training and testing sets



d. CLEVER result (boxplot with 100 repetitions) and comparison with microscopic result (black solid line), average-speed result (red dashed line), and meso result (blue dashed line)

Fig. 9. Stepwise CLEVER results for downtown Toronto.

and 74.3 tons, with a mean of 70.09 tons. Compared to average-speed and meso approaches, the improvement from CLEVER was large, leading to a similar GHG estimation result to the microscopic approach, with error less than 5%.

3.3. Application of CLEVER approach to a larger network

Based on the results of the downtown network, the CLEVER approach achieved an emission estimate that was very close to the result obtained from microsimulation. We therefore applied CLEVER to the city of Toronto.

The city network included 78% arterials, and 22% highway segments. Sample areas were selected from the city network. The sample included 173.8 km (72%) arterials and 68.8 km (28%) highways, occupying 8% of the total length of the city network. Although this percentage was smaller than the sample of roads in the downtown network, the city sample had reasonable coverage in terms of delay, density, speed and flow. The distribution of the four parameters was also similar to the full city network, as illustrated in Fig. 10.

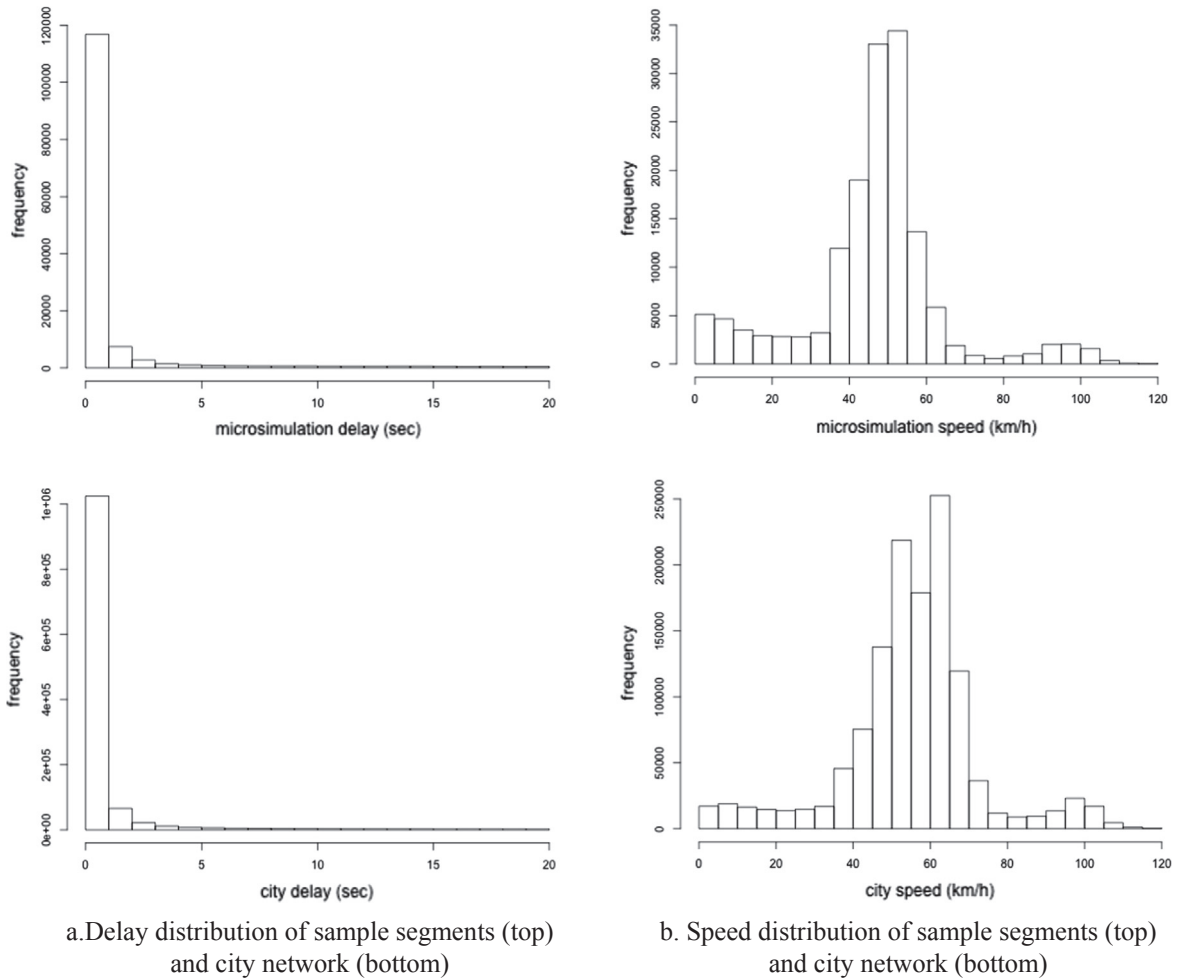


Fig. 10. Delay, speed, density and flow distribution of sample area and of full city road network.

Fig. 11 presents the results of CLEVER. Compared to the downtown network, the speed in Fig. 11a showed greater variability within clusters. In Fig. 11b, the maximum validated EF was smaller, indicating less congestion overall in the city of Toronto compared to downtown. Moreover, the EFs of each cluster were centrally distributed with fewer outliers than the downtown network. The errors of estimated testing set (micro result as baseline) of most of the clusters in Fig. 11c were centered around 0, showing that the validated EFs are representative of each cluster, with a few exceptions which were due to small numbers of segments in these clusters.

Fig. 11d shows the GHG estimation from CLEVER compared to the estimate generated based on the average-speed model. The emission estimate produced by morning peak traffic in the city of Toronto was around 585.18 tons based on the CLEVER method, whereas the result from average-speed method was 26% higher than that from CLEVER. When the event traffic conditions within the mesosimulation time interval are not constant, actual EFs will show a large variability, and the average speed based EF from the driving cycle embedded within MOVES cannot be representative of all of the events. Under this scenario, the mesoscopic approach is likely to give incorrect results, and either overestimation or underestimation can occur.

4. Discussion and conclusion

Previous studies have presented various traffic emission models, using either instantaneous speed, or average speed to estimate emissions. The challenge with microscopic models is that the data are hard to obtain, and microscopic traffic modelling is computationally intensive. Using only average speed, the calculation process is fast, while it loses detailed vehicle activities and the result is far from accurate especially under traffic congestion. Some studies have attempted to bridge the gap between the two methods by categorizing traffic conditions into predefined classes. However, the categorization is usually based on average speed and number of stops, and the variation in traffic conditions due to other factors is still neglected, therefore the predefined classes are not representative of all traffic scenarios.

In this paper, three scopes of AIMSUN traffic simulation (microscopic, mesoscopic and hybrid) were applied to downtown Toronto and total GHGs were estimated based on four approaches: microscopic model (“micro” approach), mesoscopic model with average

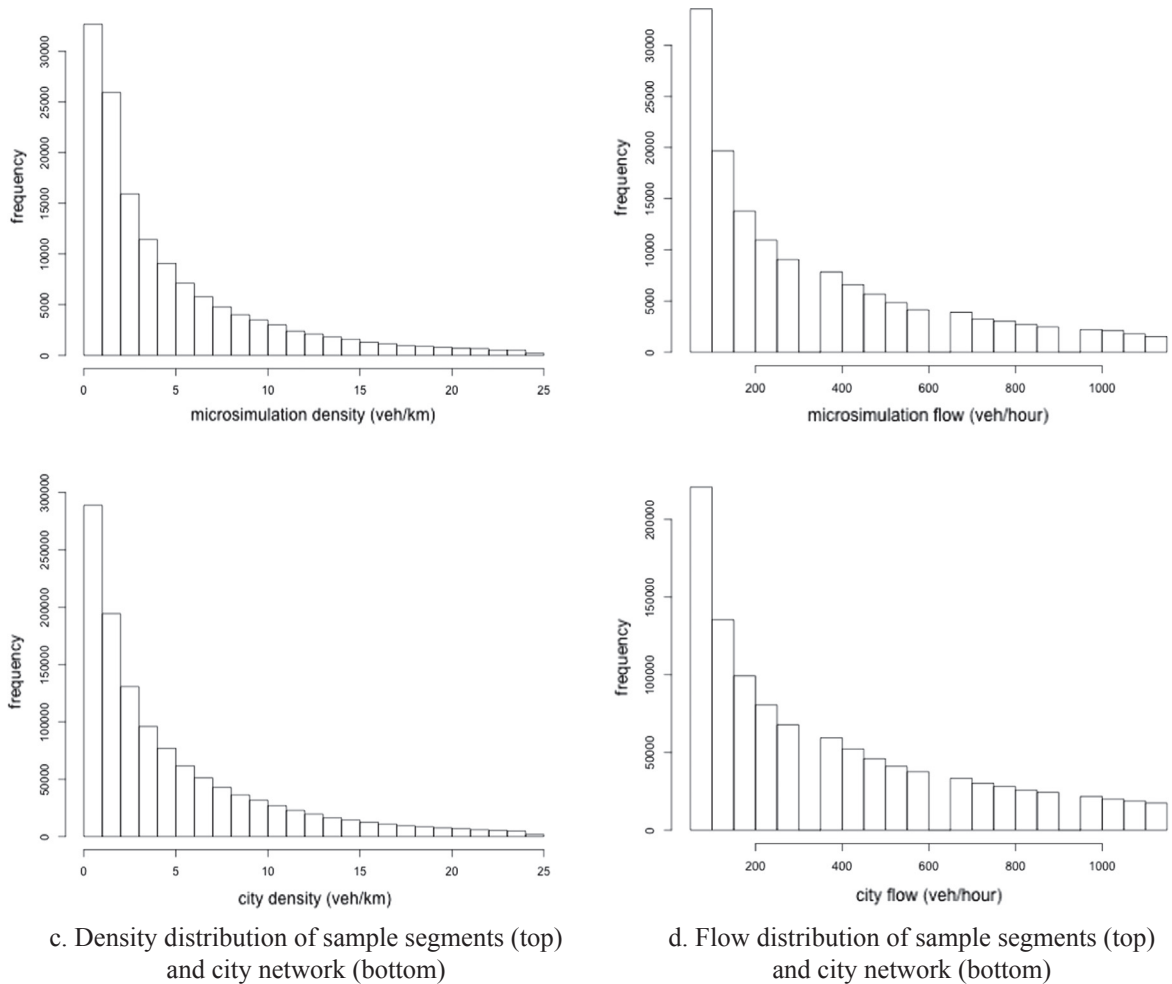


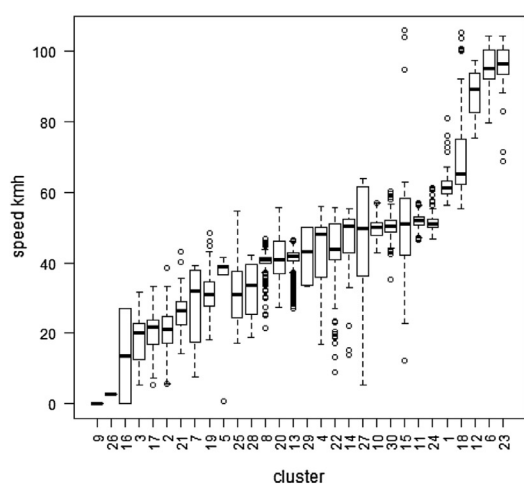
Fig. 10. (continued)

speed (“average-speed” approach), mesoscopic model with average speed and its standard deviation (“meso” approach), and hybrid approach CLEVER. Since the micro approach used a second-by-second speed profile, the emission estimate obtained was considered to be the most accurate. Based on this approach, the downtown network emitted 70.06 tons of GHGs during the morning peak period. The average-speed approach led to the largest gap with respect to the micro approach, underestimating by around 30%. Due to the inconsistent traffic conditions in the downtown network, the speed distribution cannot be assumed as normal, therefore the meso approach did not exhibit reasonable performance. Compared to the average-speed method, taking speed variation into consideration resulted in little improvement, but the emission estimate remained significantly lower than the microscopic result. Using the hourly mean and standard deviation of speed, mean and standard deviation of delay, capacity, density, and VKT as the cluster criteria, the average GHG estimate calculated by CLEVER is 70.09 tons, exhibiting the smallest difference compared to the microscopic model.

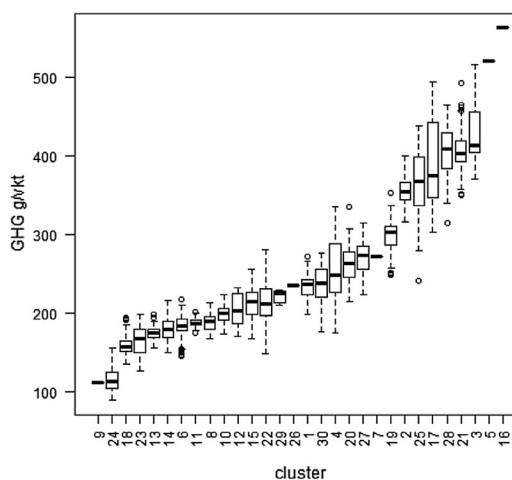
The CLEVER approach was also applied to the city of Toronto, whereby running full-network microsimulation is infeasible. For the city of Toronto, based on the CLEVER application result, the GHG estimate was 585.18 tons while the average-speed approach achieved an estimate that was higher than CLEVER by 26%.

Previous studies have investigated mesoscopic traffic emission models, using pre-defined categories of traffic conditions such as idling, cruising, accelerating, free flow, and over-saturated. The advantage of CLEVER is that it clusters road segments in an unsupervised manner, which ensures sufficient coverage of different traffic conditions, without ignoring variations within each pre-defined category. Moreover, the validation network used in previous studies consisted of either several driving cycles, or separated street links of short distances, while in this paper; CLEVER displayed a powerful application at a network-wide level.

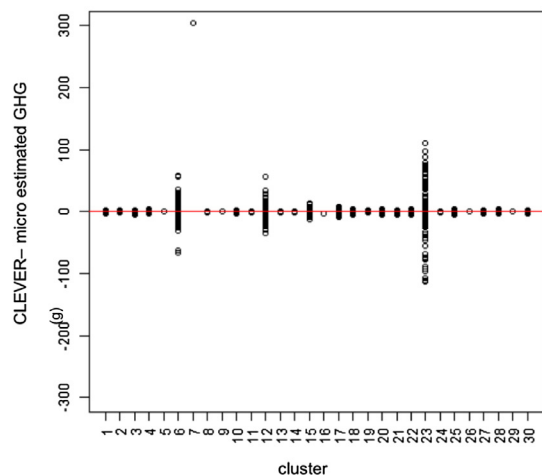
Through the comparison between CLEVER, and emission modelling approaches at different levels of resolution, this new hybrid approach demonstrates great balance between accuracy and efficiency. In this context, the CLEVER method also has the potential to be used by public agencies for large-scale traffic emission estimation. The data required in CLEVER can be split into two parts: mesoscopic road segment traffic characteristics, and instantaneous speed of sample areas. The first part can be retrieved from mesoscopic traffic simulation. AIMSUN was capable of simulating the full city network in mesoscopic mode with 20 iterations of dynamic traffic assignment in less than 3 h. The second part of this effort, which entails the generation of second-by-second vehicle



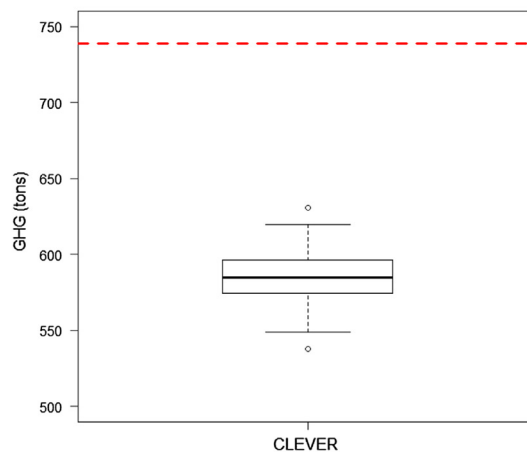
a. Speed distributions across each of 30 clusters (sorted by mean speed value of each cluster)



b. Distribution of representative EFs based on the CLEVER method; in each cluster, 100 values were generated based on the 100 random samples for the testing set (sorted by mean EF value of each cluster)



c. CLEVER minus micro emissions in the testing sets, based on 100 random samples of training and testing sets



d. CLEVER result (boxplot with 100 repetitions) and comparison with average-speed result (red dashed line)

Fig. 11. Stepwise CLEVER application results for the City of Toronto.

speeds for selected road segments, is comparatively more difficult to derive. However, CLEVER assigns a certain traffic condition with a corresponding EF to road segments. Once the relationships between these traffic condition clusters and EF sets are well established, they can be used in the same region, or regions with similar travel patterns and socioeconomic characteristics, during which the traffic demand and travel patterns do not incur significant changes.

Future research will focus on improving the algorithm of CLEVER. Additional combinations of mesoscopic traffic parameters, different numbers of clusters, as well as various clustering methods, will be tested. The cross-validation step for generating cluster EFs will also be re-structured to avoid over-fitting.

References

- Barth, M., Boriboonsomsin, K., 2008. Real-world carbon dioxide impacts of traffic congestion. *Transp. Res. Rec. J. Transp. Res. Board.* <https://doi.org/10.3141/2058-20>.
- Borge, R., de Miguel, I., de la Paz, D., Lumbreras, J., Pérez, J., Rodríguez, E., 2012. Comparison of road traffic emission models in Madrid (Spain). *Atmos. Environ.* 62, 461–471.
- Chief Corporate Officer, 2016. Toronto's 2013 Greenhouse Gas Inventory.
- Colberg, C.A., Tona, B., Stahel, W.A., Meier, M., Staehelin, J., 2005. Comparison of a road traffic emission model (HBEFA) with emissions derived from measurements in the Gubrist road tunnel. Switzerland. *Atmos. Environ.* 39, 4703–4714. <http://dx.doi.org/10.1016/J.ATMOSENV.2005.04.020>.
- Environment and Climate Change Canada, 2017. Canadian Environmental Sustainability Indicators Greenhouse Gas Emissions.
- Fujita, E.M., Campbell, D.E., Zielinska, B., Chow, J.C., Lindhjem, C.E., DenBleyker, A., Bishop, G.A., Schuchmann, B.G., Stedman, D.H., Lawson, D.R., 2012.

- Comparison of the MOVES2010a, MOBILE6. 2, and EMFAC2007 mobile source emission models with on-road traffic tunnel and remote sensing measurements. *J. Air Waste Manage. Assoc.* 62, 1134–1149.
- Gori, S., La Spada, S., Mannini, L., Nigro, M., 2013. A dynamic mesoscopic emission model for signalized intersections, in: *Intelligent Transportation Systems-(ITSC), 2013 16th International IEEE Conference on*. pp. 2212–2217.
- Hausberger, S., Rexeis, M., Zallinger, M., Luz, R., 2009. Emission Factors from the Model PHEM for the HBEFA Version 3. Rep. Nr. I-20/2009 Haus-Em 33, 679.
- Iannone, G., Guarnaccia, C., Quartieri, J., 2013. Speed distribution influence in road traffic noise prediction. *Environ. Eng. Manag. J.* 12, 493–501.
- Jamshidnejad, A., Papamichail, I., Papageorgiou, M., De Schutter, B., 2017. A mesoscopic integrated urban traffic flow-emission model. *Transp. Res. Part C Emerg. Technol.* 75, 45–83.
- Jun, J., 2010. Understanding the variability of speed distributions under mixed traffic conditions caused by holiday traffic. *Transp. Res. Part C Emerg. Technol.* 18, 599–610.
- Kamel, I., Shalaby, A., Abdulhai, B., 2018. Simulation-Based Dynamic Traffic and Transit Assignment Model for the Greater Toronto Area. In: *97th Transportation Research Board Annual Meeting*.
- Koupal, J., Cumberworth, M., Michaels, H., Beardsley, M., Brzezinski, D., 2003. Design and implementation of MOVES: EPA's new generation mobile source emission model. *Int. Emiss. Invent. Conf.* 1–15.
- Ligterink, N.E., Lange, R.D.E., 2008. Refined vehicle and driving-behaviour dependencies in the VERSIT + emission model. *Environ. Transp.* 1–8.
- Liu, H., Xu, Y.A., Rodgers, M.O., Akanser, A., Guensler, R., 2016. Improved Energy and Emissions Modeling for Project Evaluation (MOVES-Matrix).
- Ministry of Transportation Ontario, 2017. Statistics and Management Reporting.
- Park, S.S., Kozawa, K., Fruin, S., Mara, S., Hsu, Y.-K., Jakober, C., Winer, A., Herner, J., 2011. Emission factors for high-emitting vehicles based on on-road measurements of individual vehicle exhaust with a mobile measurement platform. *J. Air Waste Manage. Assoc.* 61, 1046–1056.
- Rakha, H., Ahn, K., Trani, A., 2004. Development of VT-Micro model for estimating hot stabilized light duty vehicle and truck emissions. *Transp. Res. Part D Transp. Environ.* 9, 49–74.
- Rakha, H., Ahn, K., Trani, A., 2003. Comparison of MOBILE5a, MOBILE6, VT-MICRO, and CMEM models for estimating hot-stabilized light-duty gasoline vehicle emissions. *Can. J. Civ. Eng.* 30, 1010–1021.
- Rakha, H., Yue, H., Dion, F., 2011. VT-Meso model framework for estimating hot-stabilized light-duty vehicle fuel consumption and emission rates. *Can. J. Civ. Eng.* 38, 1274–1286.
- Smit, R., Ntziachristos, L., Boulter, P., 2010. Validation of road vehicle and traffic emission models – a review and meta-analysis. *Atmos. Environ.* 44, 2943–2953.
- Smit, R., Smokers, R., Rabé, E., 2007. A new modelling approach for road traffic emissions: VERSIT+. *Transp. Res. Part D Transp. Environ.* 12, 414–422.
- TSS-Transportation Simulation Systems, 2017. Estimation of Traversal OD Matrices – Aimsun (Version 8.2) Help.
- TSS - Transport Simulation System, 2017. Simulation Process – Aimsun (Version 8.2) Help.
- TSS - Transport Simulation Systems, 2017. Hybrid Simulator – Aimsun (Version 8.2) Help.
- United States Environmental Protection Agency, 2015. Exhaust Emission Rates for Light-Duty On-road Vehicles in MOVES2014: Final Report.